Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

be determined with as great accuracy for high values of r as for low values. If this be true, it follows further that for high values of r its value should be calculated to a larger number of decimals than for low values.

Since, as noted above, a table giving the sines and cosines for each minute of arc gives $\sqrt{1-r^2}$ accurate to at least three decimal places without interpolation, such a table used without interpolation is adequate where r is calculated to the number of decimals indicated in the preceding table, that is, where r is calculated so as to give $\sqrt{1-r^2}$ accurate to three decimals.* Seldom, if ever, is it desirable to calculate the value of r to a larger number of decimals; the apparent greater accuracy would be fictitious.

THE GRAPHIC REPRESENTATION OF A FREQUENCY DISTRIBUTION

By Robert E. Chaddock, *Columbia University*

In the accompanying diagrams and the explanatory text the purpose is two-fold: (1) to present a method of exposition by which the student of elementary statistical methods will be enabled to comprehend more clearly the nature of the frequency distribution and the assumptions underlying grouped data; (2) to correct or clarify the rules for locating the median and quartiles which appear in some of the widely used elementary text-books. The student is told to add one to the number of cases and to divide by two $\frac{(n+1)}{2}$ in order to locate the mid-item in the array, the value of which is the median average. But if this rule is followed, slightly different values are secured for the median, depending on whether the calculation is made from the lower toward the higher values in the array, or from the higher toward the lower, whereas either procedure should yield exactly the same result. The reason is clear. One extra item has been inserted. After all, what we seek is not a case but a value in an array of items arranged in a regularly ascending or descending order of values, which are assumed to be evenly distributed over the scale within a given class interval so that progress in one direction shows constantly increasing values, and in the other direction constantly decreasing values.

---

* For this purpose the most convenient tables the writer has seen are in G. W. Jones: *Logarithmic Tables.* London: Macmillan, 1898. These tables give the natural and logarithmic functions in parallel columns for each minute of arc.

DIAGRAM I.

DISTRIBUTION OF PIECE-RATE EARNINGS IN 50 CENT GROUPS

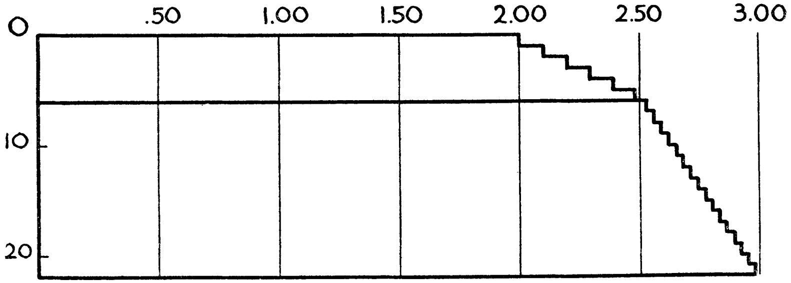


Fig. 1 SHOWING INDIVIDUAL VARIATIONS

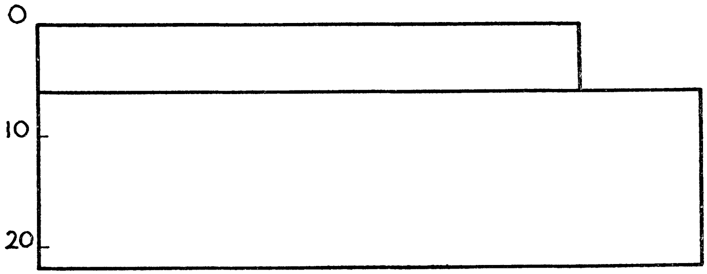


Fig. 2 SHOWING GROUP VARIATIONS

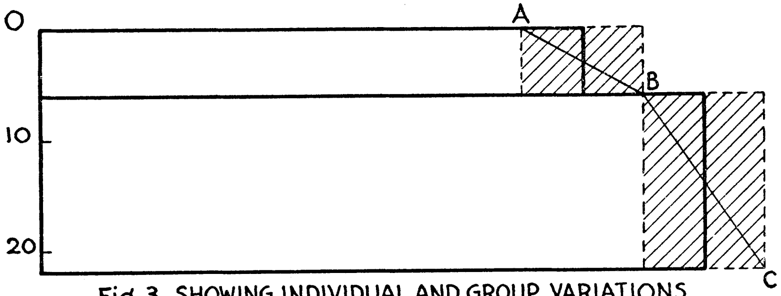


Fig. 3 SHOWING INDIVIDUAL AND GROUP VARIATIONS

The data in the illustration, presented in the form of a frequency table in Diagram II, are distributed with a fair degree of regularity because piece-rate earnings are used in 50 cent groups.* The amount of the wage is shown on the horizontal scale, and the number of wage earners on the vertical scale. Figure 1 in Diagram I shows the first two groups at the lower values on the scale distributed by individual wage earners as well as by groups, *on the assumption of even distribution* of 6 cases in the first group and 16 cases in the second group. A rectangular step on the horizontal scale and a space on the vertical scale are shown for each wage item. The distance between the limiting lines on the vertical scale shows the number of items in each group, 6 and 16 respectively.

Figure 2 shows the same facts, but the 6 items in the lower group are now given the same value on the horizontal scale concentrated at the middle of the 50 cent range, at $2.25, as would be given in fact in calculating the arithmetic average of the entire array shown in Diagram II. Likewise, the 16 items of the second group are concentrated at $2.75.

Figure 3 combines the facts of both figures 1 and 2. The cross-hatched areas show two things: (a) the range along the horizontal scale over which the items of the group are actually distributed from the smallest value to the largest, and (b) the number of items in each group, measured on the vertical scale. The diagonal *A*-*B* shows the increasing values in that group of 6 items, from the lowest values at *A* to the greatest at *B*, which was indicated by the steps in figure 1. The diagonal *B*-*C* shows the same for the 16 items of the second group, the steps in each case having been smoothed by the straight lines *A*-*B*-*C*.

Here it may be objected that values are not usually distributed over the range of class intervals so regularly as indicated in figures 1 and 3. This situation is admitted. They may be more closely grouped at the lower or upper limits of the given interval or over the upper half of the interval as compared with the lower half. But this very objection merely emphasizes the importance of the choice of the width of the interval for classifying the items, and the importance of locating experimentally the lower and upper limits of the interval. If natural massing points of the items can be discovered, these should constitute the mid-values of the class interval, shown in figure 2, as far as possible, at which all the values of the group are concentrated for practical purposes in calculating the arithmetic average. Thus, less error will be introduced in calculating the arithmetic average from grouped data as compared with the true average calculated from the individual items without grouping. In most data some classification and grouping of items is essential for ease and clearness in handling.

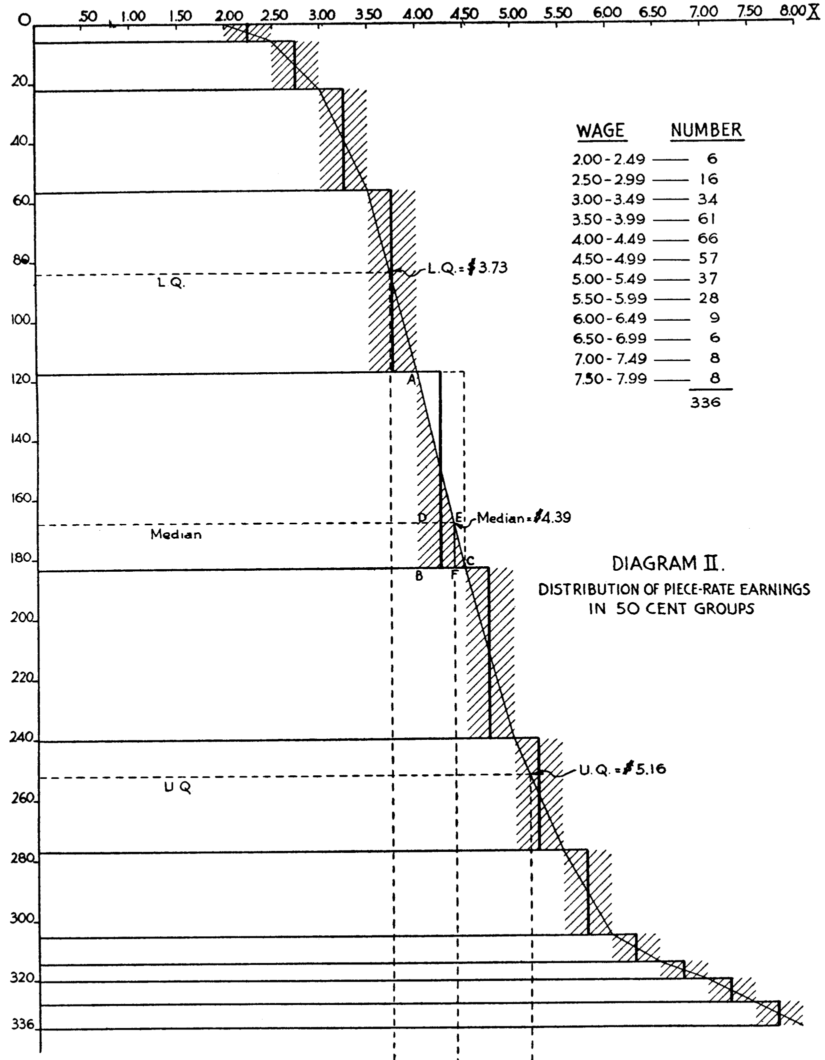
---

* The unit in which wages are paid is not capable of indefinite subdivision as in the case of measurement of height, weight, or time. This fact may affect the distribution over the class interval and the massing of cases at certain values on the scale. This is especially true in time wages, but we are using piece-rate earnings where the smallest subdivision of the unit of payment is one cent. Therefore, within any class interval of 50 cents there are many subdivisions possible, and the assumption for such data as height and weight may be made and the same methods of treatment may be employed.

In Diagram II the procedure of figure 3 in Diagram I has been carried out for each group in the entire array of 336 wage items. The limiting lines on the vertical scale show the number of cases in each group, as does also the height of the cross-hatched areas. The total height of cross-hatched areas indicates the total of the items, 336. In each group the width of the cross-hatched area is the same, measuring on the horizontal scale the range of variation for that group, the 50 cent interval. Now, if we project or let fall all the cross-hatched areas upon a common base line *O-X*, shown in Diagram III, in the exact relation and size on the horizontal and vertical scales in which they are found in Diagram II, the result will be a *surface of frequency,* sometimes called a *histogram* or *column diagram.* The height of each rectangular area above the base line represents the proper number of cases in that group, and the total area represents the total items. The total cross-hatched area in Diagram II is identical with the total area of Diagram III. By the dotted line connecting the mid-points of the tops of the columns, in Diagram III, the histogram may be readily converted into the *frequency polygon,* and by appropriate methods this may be further smoothed into a *frequency curve,* showing the distribution of values around a central tendency. This procedure is fundamental in statistical analysis.

THE EXACT LOCATION OF THE MEDIAN VALUE IN GROUPED DATA

Using Diagram II and noting that the number of items is even (336), we must locate a value on the horizontal scale, such that there will be as many value items above as below. Half the items would be $\frac{336}{2} =$ 168 items. If we count down on the vertical scale through the first four groups there will be 117 items. This brings us to the wage item of $4.00 on the horizontal scale. If we count in addition all of the fifth group, which has 66 items in it, there will be 183 items in the first five groups, up to $4.50 on the horizontal scale. But we wish a value that will show only 168 items below and the same above. Therefore, the median value must lie somewhere in the fifth group between $4.00 and less than $4.50. This group of 66 items is distributed evenly, by assumption, over the 50 cent interval. The lowest value in this group is at $4.00 and the highest just under $4.50, as shown by the diagonal *A-C* and the distance *B-C* in Diagram II. How many of the 66 items do we need,



WAGE	NUMBER
2.00 - 2.49	6
2.50 - 2.99	16
3.00 - 3.49	34
3.50 - 3.99	61
4.00 - 4.49	66
4.50 - 4.99	57
5.00 - 5.49	37
5.50 - 5.99	28
6.00 - 6.49	9
6.50 - 6.99	6
7.00 - 7.49	8
7.50 - 7.99	8
	336

DIAGRAM II.
DISTRIBUTION OF PIECE-RATE EARNINGS
IN 50 CENT GROUPS

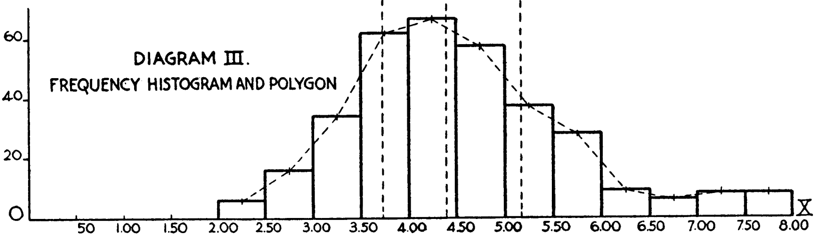


DIAGRAM III.
FREQUENCY HISTOGRAM AND POLYGON

added to the 117 of the first four groups, to complete the 168 items which must lie below the median value of the entire array? There are needed $168-117=51$ items to complete the 168 within the fifth group, as shown by the distance *A-D* measured on the vertical scale. Now these items in this group are increasing in size as we count from *A* toward *D*, as shown along the diagonal *A-C*. Beginning with $4.00 at *A*, the lower limit of the group, how much greater will the value be when we have counted 51 items along the vertical scale toward *B?* *A-B* represents 66 items, *B-C* represents 50 cents, and *A-D* represents 51 items. What will be the value range represented by *D-E?* The triangles *ABC* and *ADE* are similar. Therefore, the proportion $AB:BC::AD:DE$ will be true. The value to be found is *DE*. Substituting known values, we have 66:50 cents::51:DE, or $66DE=2550$, and $DE=39$ cents. Adding 39 cents to $4.00, the lower limit of the group, we have $4.39, the median value.

Or, reasoning in terms of proportions, on the assumption of even distribution, it follows that if 66 items are distributed over the entire range of 50 cents (*B-C*), in the fifth group, 51 items which must be added in order to have 168 items below the median value, will range over $\frac{51}{66}$ of 50 cents, or 39 cents (*D-E*), which amount must be added to the lowest wage item of the group, $4.00, in order to arrive at the median value at *E*.

Exactly the same value for the median will be secured if we begin at the bottom of Diagram II, the highest wage values, and count upward on the vertical scale of the diagram. As we proceed the wage items grow progressively smaller in each group. At the upper limit of the group of 66 items (*C*), where the value is $4.50, we have counted 153 items. It requires 15 more items in the group of 66 to fulfil the condition that 168 wage items must be greater than the median value. In the previous calculation we counted into this same group 51 items from the lower limit in order to arrive at the median. Now, we count 15 items from the upper limit to arrive at exactly the same point. $AD=51$, DB or $EF=15$, and $AD+EF=66$, the entire group. The wage values are now decreasing along the diagonal *C* to *E* where the median lies. The value to be subtracted from $4.50 may be secured as before by a proportion $AB:BC::EF:FC$, or 66:50 cents::15:*FC*, which gives 11 cents for the value of *FC*; $4.50 − .11 = $4.39; the median value, or if 66 items range over 50 cents (*B-C*), 15 items will range over $\frac{15}{66}$ of 50 cents, or 11 cents (*F-C*).

The quartiles may be secured by dividing the items by 4 and proceeding in exactly the same manner to locate the value in the array which has $\frac{1}{4}$ of the items below and $\frac{3}{4}$ of the items above it, and also the value in the array which has $\frac{3}{4}$ of the items below and $\frac{1}{4}$ above it.

Let a perpendicular fall from E in Diagram II upon the base line O-X and it will divide the area of Diagram III into two equal parts, according to the definition of the median. Like perpendicular lines, let fall from the quartile positions of Diagram II, will divide Diagram III into four equal areas.

The deciles and percentiles may be located in a similar manner.

OBSERVATIONS ON THE $\frac{n+1}{2}$ PROCEDURE FOR LOCATING THE MEDIAN

This rule is an expedient for locating the mid-case if the numbers in the array are odd, and a position between the two mid-cases if the number of cases is even. It is not a case we seek but a value. Changing values through the array are associated with movement in one direction or the other on the assumption of an even distribution over the class interval. If one case is added it will not be possible to secure the same value for the median when calculated from both the lower and upper limits of the array because the extra case introduces a slight duplication.

It is not necessary to introduce this contradiction in the procedure. In Diagram I, figure 1, a rectangular area was shown for each item, as well as for the group. A space on the vertical scale represented each item. If the number of spaces on the vertical scale, corresponding to the number of items, is even, as in our illustration in Diagram II, we count through $\frac{1}{2}$ of the spaces on the vertical scale, which brings us to a point on the line separating the two spaces representing the two mid-items. On the horizontal scale this represents a value interpolated, or the average value of the two mid-items. If the number of items is odd, in like manner we locate the mid-point on the vertical scale and find it in the middle of the space representing the mid-item of the array, with an equal number of spaces above and below it. On the horizontal scale this represents the value of the mid-item, with an equal number of items above and below it. This may be shown by supposing that the illustration in Diagram II has 337 items. Therefore, when calculating the median, quartiles, etc., for grouped data, we should use the actual number of items instead of adding one.

SHORT TERM BORROWING IN BRITISH FINANCE, 1914–1920

The accompanying charts showing the quarterly receipts of the British Treasury during the past six and a half years make it possible to compare in the case of British public finance the relative importance of different methods of raising funds during this period. The statistics upon which the charts are based have been collected by the Division of Analysis and Research of the Federal Reserve Board in connection with studies which are being made of foreign banking and finance.

The charts clearly indicate the outstanding importance of the short term Treasury bill as a source of public funds. As in the case of the United States Treasury, the British Treasury depends upon the sale of short term bills to provide cash in anticipation of receipts from taxes and longer term loans. Before the war the same method was used, but on a scale in no sense comparable with that of the war period. Treasury bills outstanding before the war did not exceed £20,000,000, whereas in recent years they have frequently exceeded the billion pound mark.

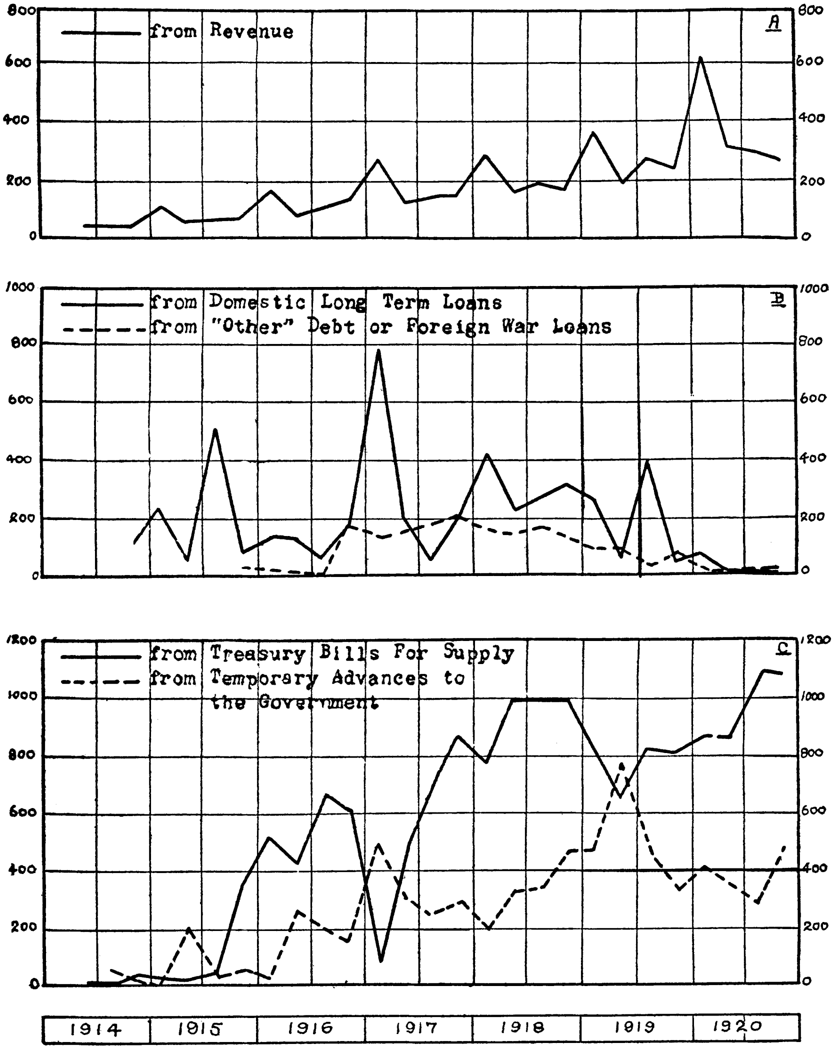
In the United States the Treasury depends entirely upon the sale of certificates of indebtedness to supply it with temporary funds. In England, on the other hand, as in France, a supplementary method is used in case Treasury bill sales are insufficient to provide the necessary cash. In France the Treasury under these circumstances obtains advances from the Bank of France. In England the Treasury applies to the Bank of England or to other public departments. The sale of Treasury bills, however, is the preferred method, since they furnish a form of investment which may be taken up by real savings, while advances from the Bank of England consist of credits created on the books of the Bank and obviously make directly for inflation. Advances from "other public departments," it is understood, are mainly drawn from the currency notes fund.

Charts B and C show that during the first year of the war, two large war loans were placed in England, providing sufficient funds in addition to those from revenue to make unnecessary the sale of Treasury bills on a large scale. During the following period, on the other hand—covering late 1915 and 1916—no long term loans were placed, and the Treasury depended very largely upon the returns from the sale of Treasury bills to provide current funds. In January 1917, the third big war loan was placed, providing for the funding of £131,000,000 of Treasury bills then outstanding; but immediately following the floating of the loan it was necessary again to turn to Treasury bills as a source of current funds. Between this time and the end of 1918, more and more Treasury bills were issued; Treasury receipts from this source amounted to £5,762,000,000, while receipts from long term loans amounted to £1,216,000,000, and from taxes £1,684,000,000. In 1919 only a small number of Treasury bills were funded in the loan floated in June, and since that time no long term loans have been placed.

The importance of the temporary advances by the Bank of England and other public departments is clearly shown in Chart C. Ever since early in 1916 it has been necessary for the Treasury to obtain such advances when receipts from Treasury bills were insufficient or when large numbers of them matured. In the first quarter of 1917 and again in 1919, these advances on ways and means actually exceeded receipts from Treasury bills. In both cases this occurred during a period when the sale of Treasury bills had been suspended (at the time of the floating of a long term loan) and when a considerable number of bills were being funded.

Although the total debt of England was somewhat reduced during 1920 (owing to repayment of certain foreign loans), the floating debt was increased by £59,000,000. The entire increase occurred in ways and means advances. How to handle this large floating debt—which amounted to £1,408,000,000 at the end of 1920—is a problem

QUARTERLY RECEIPTS BY THE BRITISH TREASURY, 1914-1920
In millions of pounds sterling



which is considerably under discussion in England. It may be possible to fund in part the Treasury bills outstanding. To pay them off even gradually will involve a heavy rate of taxation for years to come. The elimination of borrowing on ways and means is of even more pressing importance, however, since these advances result in an increase in the balances of the joint stock banks at the Bank of England.

KATHARINE SNODGRASS